

PRV

PATENT- OCH REGISTRERINGSVERKET

Patentavdelningen



A 4

Intyg Certificate

Härmed intygas att bifogade kopior överensstämmer med de handlingar som ursprungligen ingivits till Patent- och registreringsverket i nedannämnda ansökan.

This is to certify that the annexed is a true copy of the documents as originally filed with the Patent- and Registration Office in connection with the following patent application.



(71) Sökande C Technologies AB, Lund SE
Applicant (s)

(21) Patentansökningsnummer 0004144-2
Patent application number

(86) Ingivningsdatum 2000-11-13
Date of filing

Stockholm, 2001-11-12

För Patent- och registreringsverket
For the Patent- and Registration Office

Kerstin Gerdén
Kerstin Gerdén

Avgift
Fee 170:-

**CERTIFIED COPY OF
PRIORITY DOCUMENT**

AWAPATENT AB

Kontor/Handläggare

Malmö/Lars Malmqvist/LMT

C TECHNOLOGIES AB

Ansökningsnr

Vår referens

SE-2001954

Ink. t. Patent- och reg.verket

1

2000 -11- 13

TEXTPUSSLING

Huvudfoxen Kassan

Tekniskt område

Föreliggande uppfinning hänför sig till ett för-
farande och en anordning för registrering av text genom
avbildning av texten på en ljuskänslig sensor med en
5 tvådimensionell sensoryta, vilken sensor är avsedd för
digital registrering av bilder.

Teknisk bakgrund

Det finns många situationer i vilka personer önskar
sammanställa valda delar av text- eller bildinformation
10 till ett dokument som kan redigeras med lämplig mjukvara
i en dator. Ett känt sätt att mata in text- och bild-
information i en dator är att använda en scanner. Scann-
rar finns av både fast och portabel typ. En fast scanner
används för att läsa in hela sidor med text- och bild-
15 information, varvid scannern automatiskt förs över sidan
med konstant hastighet. Denna typ av scanner är inte
lämpad för inläsning av valda delar av information på en
sida. Däremot kan en portabel scanner vara lämplig för
detta ändamål.

20 I US 5 301 243 beskrivs en handhållen scanner för
läsning av tecken från en sträng av tecken på ett sub-
strat. Scannern, som förs i kontakt med substratet över
tecknen som skall läsas, har ett optiskt system som "ser"
en liten del av substratet. Det optiska systemet inne-
25 fattar en linjesensor av CCD-typ, vilken har ett flertal,
på linje anordnade, ljuskänsliga element. När scannern
förs över tecknen på substratet registreras en följd av
vertikala "skivor" av tecknen och av mellanrummen mellan
dessa. "Skivorna" lagras i scannern som en digital bit-
30 mapp. En OCR-programvara (OCR = Optical Character Recog-
nition, optisk teckenigenkänning) används därefter för
att identifiera de inlästa tecknen och lagra dem i
teckenkodad form, t ex med hjälp av ASCII-kod, i ett

2000 -11- 13

2

Huvudfoxen Kasson

minne. Teckenigenkänningen kan antingen göras i scannern eller i en extern dator till vilken de bitmappade tecknen har lästs över.

En annan typ av handhållen scanner för textinmatning
5 beskrivs i US 4 949 391. Denna scanner har till skillnad från den ovan beskrivna en tvådimensionell sensor, som registrerar bilder av den underliggande ytan då scannern förflyttas över den. Scannern kan bara förflyttas i en riktning som bestäms av ett hjul som anligger mot ytan.
10 Innan de registrerade bilderna sätts samman till en större bild tas redundant information bort ur bilderna. Den större bilden kan analyseras i en dator för att identifiera tecken.

En nackdel med de ovan beskrivna och liknande hand-
15 hållna scannrar är att deras "synfält" är förhållandevis litet. För att registrera en större informationsmängd, såsom ett textavsnitt bestående av flera rader, måste en användare därför föra scannern fram och tillbaka över ytan upprepade gånger, vilket ofta upplevs som tidsödande
20 och enformigt.

I sökandens PCT-ansökan SE99/00716 beskrivs en anordning för registrering av information från ett underlag. Anordningen kan arbeta i två moder. I den ena moden registreras ett flertal bilder på avstånd från underlaget
25 så att mer än en textrad registreras i varje bild. Ett problem med en sådan här anordning är att det krävs tidskrävande beräkningar för att pussla ihop bilderna då de är stora, om tillräcklig noggrannhet skall erhållas för att man skall kunna utföra optisk teckenigenkänning
30 på den färdiga bilden om den färdiga bilden består av ett flertal pusslade bilder. Pussling är möjlig att genomföra i en dimension, dvs att sensorn förs från vänster till höger. Om sensorn förs fram och tillbaka behövs emellertid pussling i två dimensioner, vilket ställer
35 oerhört stora krav på pusslingsberäkningarna.

2000 -11- 13

Huvudfaxen Kassan

3

Det finns således ett behov av en anordning och ett förfarande som tillåter snabbare pussling av skriven text.

Sammanfattning av uppfinningen

5 Ett ändamål med föreliggande uppfinning är att åstadkomma ett förfarande, en anordning och ett datorprogram som möjliggör snabb registrering av text i realtid.

10 Detta ändamål uppfylles med en anordning, ett förfarande och ett datorprogram enligt de vidhängande patentkraven.

En grundtanke med föreliggande uppfinning är att man registrerar ett flertal bilder av en text, vilka bilder överlappar varandra. Därefter utför man OCR (Optical
15 Character Recognition, optisk teckenigenkänning) på de registrerade bilderna och erhåller teckenuppsättningar vilka innehåller ett flertal tecken. Teckenuppsättningarna innehåller eventuellt ett antal tecken som markerar radslut om texten innehåller flera rader.
20 Därefter pusslar man ihop teckenuppsättningarna med användning av tecknen i teckenuppsättningarna. En stor fördel med det här angreppssättet är att man kan registrera relativt stora bilder åt gången utan att pusslingen av bilderna blir speciellt tidskrävande
25 eftersom man först omvandlar bilderna till tecken och sedan pusslar ihop teckenuppsättningarna. Därmed är pusslingen inte alls lika tidskrävande eftersom den effektiva upplösningen är betydligt lägre i en teckenuppsättning än vad den är i en bitmappbild. Den
30 effektiva upplösningen i teckenuppsättningen är ett tecken. Även pussling i två dimensioner blir praktiskt möjligt med uppfinningen.

Raderna i två registrerade bilder sammanfaller inte nödvändigtvis. Första raden i en första registrerad bild
35 motsvarar exempelvis tredje raden i en andra registrerad bild.

Ink. t. Patent- och reg.verket

2000-11-13

4

Huvudfaxen Kassan

Med teckenuppsättning avses i det här sammanhanget ett flertal tecken, från vilken teckenuppsättning tecknens inbördes lägen kan bestämmas. Teckenuppsättningen kan exempelvis vara en teckensträng som

5 innehåller tecken för mellanslag och radslut.

En anordning för registrering av en text enligt uppfinningen gör det genom avbildning av texten på en ljuskänslig sensor med en tvådimensionell sensoryta, vilken sensor är avsedd för digital registrering av bilder av

10 texten, vilka bilder har delvis överlappande innehåll. Anordningen utmärkes av att den är anordnad att omvandla åtminstone två av de med sensorn registrerade bilderna till vars en teckenuppsättning, innefattande ett flertal tecken, medelst teckenigenkänning. Anordningen är

15 anordnad att därefter pussla ihop teckenuppsättningarna med ledning av tecknen i teckenuppsättningarna.

Såsom nämnts ovan finns det stora fördelar med en sådan här anordning. Genom att man genomför teckenigenkänning innan man pusslar ihop bilderna undviker man

20 den beräkningskrävande operationen att pussla ihop stora bilder på bildpunktsnivå. Det finns heller ingen risk för att man inte skulle känna igen något tecken pga dålig pussling vilket är fallet när man först pusslar ihop stora digitala bilder och därefter utför teckenigen-

25 känning på tecken som möjligen ligger i ett område som överlappas av båda de bilder som skall pusslas ihop. I det senare fallet finns risken för att tecknet blir förvrängt om pusslingen inte görs med tillräckligt stor noggrannhet vilket i sin tur leder till att tecknet inte

30 känns igen vid teckenigenkänningen. Med en anordning enligt uppfinningen är risken mindre för att ett tecken inte skall kännas igen eftersom man har den ursprungliga kvaliteten på bilden när man utför teckenigenkänningen. Genom att man först omvandlar bilderna till

35 teckenuppsättningar så går själva pusslingen av teckenuppsättningarna relativt snabbt att genomföra eftersom antalet tecken är betydligt mindre än antalet

2000 -11- 13

Huvudfaxen Kassan

5

bildpunkter i de registrerade bilderna. Genom att man OCR-tolkar varje bild innan pusslingen får man flera OCR-tolkningar av samma tecken, en för varje bild där tecknet ingår, och kan välja den tolkning som givit säkrast

5 resultat.

Även om det är mest fördelaktigt att pussla ihop teckenuppsättningarna genom att jämföra de enskilda tecknen i teckenuppsättningarna är det möjligt att anordningen är anordnad att pussla ihop teckenuppsättningarna
10 med hjälp av ord i teckenuppsättningen. Således jämförs hela ord i den ena av de teckenuppsättningar som skall pusslas ihop med ord i den andra av de teckenuppsättningar som skall pusslas ihop. När man pusslar ihop ord måste man i praktiken ändå jämföra varje enskilt
15 tecken.

Med ord menas i det här sammanhanget ett flertal tecken som är avgränsade med speciella förutbestämda tecken. De speciella tecknen är exempelvis mellanslag, punkt, komma eller radslut.

20 Med fördel är anordningen anordnad att hitta textraderna i de registrerade bilderna, att hitta start och slut för ord längs textraderna, och att välja vilka av de registrerade bilderna som skall omvandlas till teckenuppsättningar med ledning av ordens identifierade start
25 och slut i de registrerade bilderna, så att endast bilder med nödvändig information omvandlas till teckenuppsättningar. Genom att identifiera start och slut för orden längs textraderna blir det möjligt att göra en grov pussling av bilderna utan att genomföra optisk teckenigenkän-
30 ning. När man har identifierat ordens start och slut korrigerar man de registrerade bilderna för rotation och perspektiv. Genom att göra en grov pussling av bilderna blir det möjligt att ta reda på hur bilderna överlappar varandra. De bilder som endast innehåller information som
35 finns fullständigt i andra bilder behöver då inte omvandlas till teckenuppsättningar och därigenom minimerar

man beräkningsarbetet för att omvandla bilderna till teckenuppsättningar.

Med fördel identifieras start och slut för ord längs
textraderna genom att man på varje punkt längs en linje
5 genom textraderna söker efter antalet mörka punkter ett
förutbestämt antal bildpunkter upp och ner från
textlinjen, varvid ett ordslut definieras som att man
inte har några mörka bildpunkter på ett förutbestämt
antal bildpunkter ovanför och nedanför textlinjen, dvs
10 man har mellanslag. För att hantera kursiv stil kan man
alternativt leta utmed en sned linje. Givetvis är det
möjligt att man istället för mörka bildpunkter letar
efter ljusa bildpunkter i fallet att texten skulle vara
ljusare än bakgrunden.

15 Med fördel omvandlas bilderna till binära bilder,
dvs bilder som endast innehåller svart och vitt, eftersom
detta underlättar den fortsatta bildbehandlingen.

Med fördel är anordningen anordnad att hitta
textraderna i de registrerade bilderna med hjälp av
20 Hough-transformering av de registrerade bilderna.
Givetvis kan man hitta textraderna på något annat sätt
också men Hough-transformering är ett effektivt sätt att
göra det på. Hough-transformering kan i korthet beskrivas
på följande sätt. Det finns ett oändligt antal rät
25 linjer som går igenom en punkt i ett XY-plan. Den rät
linjens ekvation för var och en av de här linjerna kan
uttryckas med två parametrar. Om man plottar de olika
linjernas parametrar i ett diagram erhålls en kurva som
motsvarar Hough-transformen för punkten. Man kan på det
30 här sättet plotta kurvor för vilken som helst av
punkterna i XY-planet, vilka kurvor är punkternas Hough-
transform. Två olika punkters Hough-transform kommer att
skära varandra i en punkt vilken punkt motsvarar den rät
linjens ekvation för den linje som går igenom de båda
35 punkterna. Om man på motsvarande sätt plottar Hough-
transformerna för alla de mörka punkterna i en registre-
rad bild erhålls ett stort antal skärningspunkter mellan

2000 -11- 13

Huvudfaxen Kassan

7

de olika Hough-transformerna. Man kommer emellertid att få ett maximalt antal skärningspunkter för linjer som följer textraderna.

Företehrådesvis innefattar anordningen ett minne och
5 är anordnad att lagra de registrerade bilder, som skall omvandlas till teckenuppsättningar, i minnet, och att omvandla de lagrade bilderna till teckenuppsättningar efter att registreringen av bilderna avslutats. Genom att endast lagra de registrerade bilder som skall omvandlas
10 till teckenuppsättningar minimeras minnesutrymmet som krävs för att lagra registrerade bilder. Genom att omvandla de lagrade bilderna till teckenuppsättningar efter att registreringen av bilder avslutats behöver man inte ställa höga krav på den optiska teckenigenkännings-
15 hastighet vilket vore fallet om den skulle utföras samtidigt som bilder registreras.

Anordningen är med fördel utformad så att en användare kan hålla den i handen och på avstånd från ett underlag registrera text på underlaget eftersom det är
20 enklare att hantera en handhållen anordning.

Med fördel är anordningen anordnad att korrigera bilderna med avseende på rotation innan de omvandlas till teckenuppsättningar. I fallet att man har identifierat textraderna i de registrerade bilderna kan korrigeringen
25 med avseende på rotation göras relativt enkelt. Det är en stor fördel att man har korrigerat bilderna med avseende på rotation så att detta inte behöver göras av teckenigenkänningsalgoritmen.

På motsvarande sätt är det fördelaktigt om anordningen är anordnad att korrigera bilderna med avseende på perspektiv innan de omvandlas till teckenuppsättningar eftersom den optiska teckenigenkänningen underlättas betydligt om alla bokstäver har samma storlek i de bilder på vilka teckenigenkänningen skall utföras.

35 De här korrigeringarna är speciellt fördelaktig i fallet att anordningen är utformad som en handhållen enhet som på avstånd från underlaget registrerar text på

2000 -11- 13

underlaget eftersom en användare av en sådan handhållen enhet med svårighet kan undvika att vrida och luta enheten under registreringen av bilderna.

Givetvis är det möjligt att anordningen är utformad på något annat sätt än som en helt och hållet handhållen enhet. Exempelvis kan anordningen vara utformad som ett läshuvud som är kopplat till en beräkningsenhet i vilken de registrerade bilderna behandlas. Anordningen består enligt den här utföringsformen av två fysiskt åtskilda enheter.

Anordningen är med fördel anordnad att pussla ihop teckenuppsättningarna genom att jämföra teckenuppsättningarna parvis, varvid teckenuppsättningarna jämförs i ett antal inbördes lägen förskjutna i förhållande till varandra. Således jämförs ett första tecken i en första teckenuppsättning med tecken i den andra teckenuppsättningen tills överensstämmelse uppnås eller tills det första tecknet i den första teckenuppsättningen jämförts med alla tecknen i den andra teckenuppsättningen. Därefter jämförs överensstämmelsen för de andra tecknen i den första teckenuppsättningen med tecknen i den andra teckenuppsättningen. Genom att göra jämförelsen för ett stort antal olika inbördes lägen kan ett flertal poängsummor erhållas där poängsumman återspeglar överensstämmelsen mellan de båda teckenuppsättningarna för det specifika läget. På det här sättet kan ett optimalt inbördes läge erhållas.

Företrädesvis är anordningen anordnad att lagra de registrerade bilderna som skall omvandlas till teckenuppsättningar tillsammans med ett ordningsnummer som anger i vilken ordning bilderna registrerats och att pussla ihop teckenuppsättningarna med ledning av ordningsnumret för bilderna som motsvarar teckenuppsättningarna. Speciellt i fallet att man först registrerar alla bilder och först därefter påbörjar teckenigenkänningen och pusslingen är det fördelaktigt att använda sig av ordningsnummer för de registrerade bilderna

- eftersom man då har ett stort antal bilder som skall pusslas ihop. Genom att pussla bilderna med ledning av ordningsnumret kan ett enklare datorprogram användas för pusslingen än vad som är fallet om man inte använder sig
- 5 av ordningsnummer utan pusslar bilderna efter bästa överensstämmelse. I det senare fallet måste en av bilderna jämföras med alla de andra bilderna för att se vilken som passar bäst och därefter måste det här förfarandet återupprepas för resten av bilderna.
- 10 Enligt en andra aspekt av föreliggande uppfinning tillhandahålles ett förfarande för registrering av text på ett underlag innefattande steget att avbilda och digitalt registrera bilder av texten, vilka bilder har delvis överlappande innehåll. Förfarandet utmärkes av att
- 15 det även innefattar stegen att omvandla åtminstone två av de registrerade bilderna till vars en teckenuppsättning, med ett flertal tecken, medelst optisk teckenigenkänning, och att pussla ihop teckenuppsättningarna med ledning av tecknen i teckenuppsättningarna.
- 20 Det är fördraget att det med sensorn registrerade området är så stort att ett flertal textrader avbildas i en registrerad bild.
- De olika särdrag som beskrivits i samband med anordningen ovan kan implementeras även i förfarandet enligt
- 25 uppfinningen med samma fördelar som nämnts ovan.
- Enligt en tredje aspekt av föreliggande uppfinning tillhandahålles ett minnesmedium, vilket kan avläsas av en dator och på vilket är lagrat ett datorprogram som är
- 30 avsett att användas för omvandling av digitala bilder, som registrerats med en bildsensor, till text. Minnesmediet utmärkes av att datorprogrammet innefattar instruktioner för att bringa datorn att mottaga digitala bilder som insignaler, att omvandla de digitala bilderna till teckenuppsättningar, med ett flertal tecken, medelst
- 35 teckenigenkänning, och att pussla ihop teckenuppsättningarna med ledning av tecknen i teckenuppsättningarna.

2000-11-13

Huvudfoxen Kassan

10

De olika särdrag som beskrivits i samband med anordningen ovan kan implementeras även i minnesmediet enligt uppfinningen med samma fördelar som nämnts ovan.

Givetvis kan de olika särdragen som beskrivits ovan
5 kombineras i samma utföringsform.

För att ytterligare belysa uppfinningen kommer i det följande detaljerade utföringsformer av uppfinningen att beskrivas, utan att emellertid uppfinningen skall anses vara begränsad härtill.

10 Kort beskrivning av ritningarna

Fig 1 visar schematiskt en anordning enligt en föredragen utföringsform av föreliggande uppfinning.

Fig 2 visar schematiskt hur olika delar i anordningen är kopplade till varandra.

15 Fig 3 åskådliggör hur bilder registreras av text på ett papper enligt en föredragen utföringsform av föreliggande uppfinning.

Fig 4 åskådliggör principen för Hough-transformation.

20 Fig 5 visar maxpunkter för Hough-transformen av två olika bilder.

Fig 6 visar hur man detekterar ordsslut och början på ord enligt en föredragen utföringsform av föreliggande uppfinning.

25 Fig 7 åskådliggör hur man kan använda de detekterade slutet och början på ord för att grovt sortera bort de bilder som inte behöver omvandlas till teckenuppsättningar.

Fig 8 visar två teckenuppsättningar som skall
30 pusslas ihop.

Fig 9 åskådliggör textrader i en registrerad bild.

Fig 10 visar ett flödesschema över funktionen hos ett datorprogram enligt en föredragen utföringsform av föreliggande uppfinning.

35 Beskrivning av föredragna utföringsformer

I den i fig 1 visade utföringsformen av anordningen enligt uppfinningen har denna ett hölje 1 med ungefär

2000 -11- 1 3

Huvudfoxen Kassan

11

samma form som en konventionell överstrykningspenna.

Höljets ena kortände har en öppning 2, vilken är avsedd att riktas mot ett område på ett underlag i form av en informationsbärare som en användare önskar avbilda.

- 5 Informationsbäraren kan exempelvis utgöras av ett papper eller någon annan typ av lagringsmedium, från vilket man önskar registrera text i teckenformat.

Höljet 1 inrymmer i huvudsak en optikdel 3, en elektronikdel 4 och en strömförsörjningsdel 5.

- 10 Optikdelen 3 innefattar dels ett linssystem 7, dels lysdioder 6, dels en ljuskänslig sensor 8 som utgör gränssnitt mot elektronikdelen 4. Lysdioderna 6 kan användas för att öka belysningen i förhållande till den befintliga belysningen.

- 15 Linssystemets 7 uppgift är att avbilda ett område på informationsbäraren på den ljuskänsliga sensorn 8 på ett så korrekt sätt som möjligt. Den ljuskänsliga sensorn 8 utgörs t ex av en tvådimensionell, CCD-enhet (CCD = Charge Coupled Device, laddningskopplad anordning)
- 20 med inbyggd A/D-omvandlare. Sådana sensorer är kommersiellt tillgängliga. Sensorn 8 är här monterad på ett eget kretskort 11. Strömförsörjningen till anordningen erhålls från ett batteri 12 som är monterat i ett separat fack 13 i höljet. Givetvis är det möjligt att använda även andra
- 25 typer av sensorer, vilka omvandlar ljus till digitala signaler.

- I blockschemat i fig 2 visas elektronikdelen 4 schematiskt. Elektronikdelen 4 innefattar en processor 20, vilken via en buss 21 är kopplad till ett läsminne 22, i
- 30 vilket processorns program är lagrat, till ett skriv och läsminne 23, vilket utgör processorns arbetsminne och i vilket bilderna från sensorn liksom tecken som tolkas från de registrerade bilderna lagras, till en styrlogikenhet 24 och till sensorn 8.

- 35 Styrlogikenheten 24 är kopplad till ett antal periferienhet, som innefattar en display 25, som är monterad i höljet, en IR-sändtagare 26 för överföring av informa-

2000-11-13

Huvudfoxen Kassan

12

tion till/från en extern dator, knappar 27 medelst vilka användaren kan styra anordningen, samt en driftindikeringsanordning 28 bestående av t ex en andra uppsättning lysdioder, som indikerar om registrering sker eller ej. I 5 styrlogikenheten 24 genereras styrsignaler till minnena, sensorn och periferienheterna. Styrlogiken hanterar även generering och prioritering av avbrott till processorn. Knapparna 27, IR-sändtagaren 26, displayen 25 och lysdioderna 6 nås genom att processorn skriver och läser i 10 registret i styrlogikenheten. Knapparna 27 genererar avbrott till processorn 20 när de aktiveras.

Anordningens funktion kommer nu att beskrivas med hänvisning till figurerna. I fig 3 visas ett papper 9 med ett flertal textrader 10. När en användare aktiverar 15 anordningen medelst knapparna 27 och för den över pappret med öppningen 2 riktad mot pappret registreras fyra bilder 14, 15, 16, 17. Var och en av bilderna 14, 15, 16, 17 omvandlas till teckenuppsättningar såsom de som visas i fig 8. Därefter pusslas teckenuppsättningarna ihop så 20 att en hel text erhålles. Såsom visas i fig 3 är en första bild 14 vriden i förhållande till en andra bild 15 vilken i sin tur är vriden i förhållande till en tredje bild 16. För att optimera den optiska teckenigenkänningen i de registrerade bilderna 14, 15, 16, 17 är det 25 fördelaktigt att man känner till textradernas orientering i bilden. Därför detekteras textradernas orientering innan teckenigenkänningen utförs. Detekteringen av textradernas orientering görs med fördel med hjälp av Hough-transformering.

30 Med hänvisning till fig 4 och 5 kommer nu Hough-transformering att beskrivas översiktligt. I fig 4A visas två punkter i planet med koordinaterna X och Y. I fig 4B visas Hough-transformen av punkterna i fig 4A. En första punkt 18 har en första Hough-transform 19 vilken 35 beskriver alla de rätta linjer som går igenom den första punkten 18 i fig 4A som funktion av parametrarna θ och ρ där θ är vinkeln för den rätta linjen genom den första

Ink. t. Patent- och reg.verket

2000-11-13

13

Huvudfaxen Kossan

punkten 18 och p är den råta linjens avstånd från origo i XY-planet. På motsvarande sätt har den andra punkten 29 en andra Hough-transform 30. Den första och den andra Hough-transformen i fig 4B skär varandra i punkten 31 och
5 den råta linje som motsvarar av punkten 31 i fig 4B motsvarar den råta linje som går genom både den första punkten 18 och den andra punkten 29 i fig 4A.

I fig 9 visas den andra registrerade bilden 15 från fig 3. Den registrerade bilden 15 består av ett flertal
10 bildpunkter i vilka textrader skall hittas. För varje svart punkt i den andra registrerade bilden 15 görs en Hough-transform. Hough-transformerna läggs in i samma diagram. I fig 9 visas en första linje 32 och en andra linje 33. Linjer som har ungefär samma riktning som den
15 första linjen 32 kommer att skära ett större antal punkter än linjer som har ungefär samma lutning som den andra linjen 33 eftersom de mörka bildpunkterna i den registrerade bilden 15 ligger längs textrader som har ungefär samma riktning som den första linjen 32.

I fig 5 åskådliggörs hur maxima för antalet skärningspunkter kan användas för att bestämma textradernas orientering. Cirkelarna 73 motsvarar de punkter i Hough-transformdiagrammet där flest linjer skär varandra. Cirkelarna 73 ligger på en rät korsningslinje 34 och mot-
20 svarar alla samma registrerade bild. De olika cirkelarna 73 motsvarar de olika raderna i den registrerade bilden. Från avståndet mellan cirkelarna 73 kan avståndet mellan textraderna bestämmas. Korsningslinjens 34 läge längs θ -axeln indikerar hur den registrerade bilden är roterad.
25 Den andra korsningslinjen 35 i fig 5 går igenom ett flertal andra maxima som indikeras med kryss 36 i diagrammet. Den andra korsningslinjens 35 lutning indikerar att bilden har perspektiv dvs textraderna ser ut att gå ihop i den registrerade bilden. Den andra
30 korsningslinjens 35 förskjutning i förhållande till den första linjen längs θ -axeln indikerar att textraderna är roterade i den registrerade bilden.

2000 -11- 1 3

Huvudfaxen Kassan

14

- Efter att textraderna har identifierats görs en identifiering av ordens början och slut i den registrerade bilden. I fig 6 indikeras hur bokstaven "e" 37 detekteras. Såsom visas i fig 6 räknas antalet mörka bildpunkter 7 bildpunkter i höjddled vinkelträt mot linjen 38. Antalet mörka pixlar är 0 fram till början 39 på bokstaven "e" och blir återigen 0 vid slutet 40 av bokstaven "e". Då antalet mörka pixlar varit 0 under en förutbestämd period detekteras detta som ett ordslut.
- 10 Med hänvisning till fig 7 visas hur orden indikeras som rektanglar varvid rektanglarnas början 41 indikerar början på ett ord och rektanglarnas slut 42 indikerar slutet på ett ord. Fig 7A motsvarar en första registrerad bild och fig 7B motsvarar en andra registrerad bild. Ett 15 första ord 43 i fig 7A har sin motsvarighet i ett andra ord 44 i fig 7B. På motsvarande sätt har ett tredje ord 45 i fig 7A sin motsvarighet i ett fjärde ord 46 i fig 7B. Genom att matcha de båda bilderna kan man ta reda på hur de registrerade bilderna överlappar varandra.
- 20 Genom att använda ordens längd för varje rad är det möjligt att genomföra en grovpussling av de två bilderna. Bilderna grovpusslas så att en följd av ordlängder i den första registrerade bilden överensstämmer med en följd av ordlängder i den andra registrerade bilden. Ord­längderna 25 längs olika rader i olika rader i den första registrerade bilden skall således överensstämma med ordlängder längs motsvarande rader i den andra registrerade bilden. På detta sätt kan man bestämma hur bilderna är förskjutna relativt varandra.
- 30 I fig 3 visas hur detta kan utnyttjas för att sortera bort onödiga bilder så att optisk teckenigenkänning inte behöver utföras på alla bilder som registreras. En första bild 47 och en andra bild 49 överlappar helt det område som täcks av en tredje registrerad bild 48 som 35 indikeras med streckade linjer. Genom att använda för­farandet som beskrevs i samband med fig 7 kan den tredje

2000 -11- 13

Huvudfaxen Kassan

15

registrerade bilden 48 helt lämnas därhån utan att man genomför optisk teckenigenkänning på den.

I fig 8 visas en första teckenuppsättning 50 med ett flertal tecken 60 som motsvarar den första registrerade bilden 47 i fig 3 och en andra teckenuppsättning 51 med ett flertal tecken 61 som motsvarar den andra registrerade bilden 49 i fig 3. Texten i den första teckenuppsättningen och den andra teckenuppsättningen följer textradernas riktning 62. Den första teckenuppsättningen 50 och den andra teckenuppsättningen 51 pusslas ihop genom att man jämför tecken i de båda teckenuppsättningarna. Sålunda börjar man med att jämföra det första tecknet 52 i den första teckenuppsättningen med vart och ett av tecknen i den andra teckenuppsättningen 51. Därefter gör man på motsvarande sätt för det andra tecknet 63 och det tredje tecknet 64 i den första teckenuppsättningen 50. God överensstämmelse erhålls när tecknen i ordet "skilled" 53 jämföres med ordet "skilled" 54 på den första raden i den andra teckenuppsättningen 51. Eftersom ett ord kan förekomma på många ställen i en text så utgår man från den först hittade överensstämmelsen och jämför därefter resten av texten för det här läget varvid en total poängsumma erhålls som indikerar hur väl de båda teckenuppsättningarna 25 överensstämmer för det här läget. Därefter återupprepas det här steget för nästa läge där överensstämmelse erhålls. Slutligen väljs det läge vars poängsumma indikerar bäst överensstämmelse. I fig 8 är texten på engelska men en fackman inom området inser att texten 30 lika gärna kunde vara på något annat språk och att texten i fig 8 endast används för att åskådliggöra funktionen hos anordningen.

Genom att bilder pusslas först med grovpussling och sedan med pussling på teckennivå blir det möjligt att 35 genomföra pussling i två dimensioner utan att alltför stor datorkapacitet erfordras. Om läsanordningen först förs åt höger som i fig 3 och sedan nedåt och därefter åt

2000 -11- 13

Huvudfaxen Kassan

16

vänster erhålls en större yta som skall pusslas i två dimensioner. Genom att bestämma delbildernas inbördes relation med grovpussling och sedan på teckennivå är det relativt enkelt att åstadkomma både horisontell och

- 5 vertikal pussling. Genom egenskapen att raderna bestämmer den vertikala positionen med stor noggrannhet erhålls inriktningensmöjlighet. Givetvis är det inte alltid nödvändigt att först genomföra grovpussling.

- I föreliggande uppfinning finns kravet att
10 informationen befinner sig utmed identifierbara rader. Om anordningen förs över ett fotografi eller annan yta som inte är uppdelad i rader, kan detta enkelt registreras av processorn i anordningen och denna del av bilden överhoppas.

- 15 I fig 10 visas ett flödesschema över funktionen hos ett datorprogram enligt uppfinningen. Datorprogrammet är anordnat att exekveras i elektronikdelen 4 i anordningen. I ett första steg 55 mottages digitala bilder från sensorn 8. I ett andra steg 56 omvandlas de digitala
20 bilderna till teckensträngar medelst teckenigenkänning. I ett tredje steg 57 pusslas teckensträngarna ihop.

- En fackman inom området inser att uppfinningen ej är begränsad till de här visade utföringsformerna utan att många modifieringar är möjliga inom ramen för uppfin-
25 ningen. Givetvis behöver man inte utföra den i samband med fig 7 beskrivna grova passningen av de registrerade bilderna utan optisk teckenigenkänning kan utföras direkt på de registrerade bilderna. På samma sätt behöver inte textradernas orientering identifieras i de registrerade
30 bilderna om en optisk teckenigenkänningsalgoritm används som kan identifiera tecken även då textraderna är roterade.

2000-11-13

Huvudfaxen Kassan

17

PATENTKRAV

1. Anordning för registrering av en text genom
avbildning av texten på en ljuskänslig sensor (8) med en
5 tvådimensionell sensoryta, vilken sensor (8) är avsedd
för digital registrering av bilder (14-17), k ä n n e -
t e c k n a d av att den är anordnad
att med sensorn registrera åtminstone två bilder
(14-17) av texten, vilka bilder har delvis överlappande
10 innehåll,
att omvandla var och en av åtminstone två av de med
sensorn registrerade bilderna till en teckenuppsättning
(50, 51), innefattande ett flertal tecken (60, 61),
medelst teckenigenkänning, och
15 att därefter pussla ihop teckenuppsättningarna (50,
51) med ledning av tecknen i teckenuppsättningarna.
2. Anordning enligt patentkrav 1, vilken är anordnad
att identifiera textradernas riktning (62) i de
registrerade bilderna (14-17, 47-49),
20 att identifiera start och slut för ord längs
textraderna,
att välja vilka av de registrerade bilderna (14-17,
47-49) som skall omvandlas till teckenuppsättningar (50,
51) med ledning av ordens (53, 54) identifierade start
25 och slut i de registrerade bilderna, så att endast bilder
med nödvändig information omvandlas till teckenupp-
sättningar.
3. Anordning enligt patentkrav 2, vilken är anordnad
att hitta textraderna (62) i de registrerade bilderna med
30 hjälp av Hough-transformering av de registrerade bilderna
(14-17, 47-49).
4. Anordning enligt något av föregående patentkrav,
vilken innefattar ett minne och vilken är anordnad
att lagra de registrerade bilder (14-17, 47-49) som
35 skall omvandlas till teckenuppsättningar (50, 51) i
minnet, och

2000 -11- 1 3

att omvandla de lagrade bilderna till teckenuppsättningar (50, 51) efter att registreringen av bilder avslutats.

5. Anordning enligt något av föregående patentkrav, vilken är anordnad att korrigera de registrerade bilderna (14-17, 47-49) med avseende på rotation innan de omvandlas till teckenuppsättningar.

6. Anordning enligt något av föregående patentkrav, vilken är anordnad att korrigera de registrerade bilderna (14-17, 47-49) med avseende på perspektiv innan de omvandlas till teckenuppsättningar.

7. Anordning enligt något av föregående patentkrav, vilken är anordnad att pussla ihop teckenuppsättningarna (50, 51) genom att jämföra teckenuppsättningarna (50, 51) parvis, varvid teckenuppsättningarna (50, 51) jämförs i ett flertal inbördes lågen förskjutna (53, 54) i förhållande till varandra.

8. Anordning enligt något av föregående patentkrav, vilken är anordnad att lagra de registrerade bilder som skall omvandlas till teckenuppsättningar (50, 51) tillsammans med ett ordningsnummer som anger i vilken ordning de registrerade bilderna (14-17, 47-49) registrerats och vilken är anordnad att pussla ihop teckenuppsättningarna med ledning av ordningsnumret för de registrerade bilderna (14-17, 47-49) som motsvarar teckenuppsättningarna (50, 51).

9. Förfarande för registrering av text innefattande steget att avbilda och digitalt registrera bilder (14-17, 47-49) av texten, vilka bilder har delvis överlappande innehåll, k å n n e t e c k n a t av att det även innefattar stegen

att omvandla åtminstone två av de registrerade bilderna (14-17, 47-49) till vars en teckenuppsättning (50, 51), med ett flertal tecken, medelst teckenigenkänning, och

2021
2000 -11- 1 3

Huvudfoxen Kossan

19

att därefter pussla ihop teckenuppsättningarna (50, 51) med ledning av tecknen i teckenuppsättningarna (50, 51).

5 10. Förfarande enligt patentkrav 9, vilket även innefattar stegen

att identifiera textradernas riktning (62) i de registrerade bilderna (14-17, 47-49),

att hitta start och slut för ord (53, 54) längs textraderna,

10 att välja vilka av de registrerade bilderna som skall omvandlas till teckenuppsättningar (50, 51) med ledning av ordens (53, 54) identifierade start och slut i de registrerade bilderna (14-17, 47-49), så att endast bilder med nödvändig information omvandlas.

15 11. Förfarande enligt patentkrav 10, varvid textraderna identifieras med hjälp av Hough-transformering av de registrerade bilderna (14-17, 47-49).

20 12. Förfarande enligt något av patentkraven 9-11, varvid de registrerade bilderna (14-17, 47-49) korrigeras med avseende på rotation innan de omvandlas till teckenuppsättningar (50, 51).

25 13. Förfarande enligt något av patentkraven 9-11, varvid de registrerade bilderna korrigeras med avseende på perspektiv innan de omvandlas till teckenuppsättningar (50, 51).

30 14. Minnesmedium, vilket kan avläsas av en dator och på vilket är lagrat ett datorprogram som är avsett att användas för omvandling av digitala bilder (14-17, 47-49), till text, k ä n n e t e c k n a t av att datorprogrammet innefattar instruktioner för att bringa datorn

att mottaga digitala bilder (14-17, 47-49) som insignaler,

35 att omvandla de digitala bilderna (14-17, 47-49) till teckenuppsättningar (50, 51), med ett flertal tecken (60, 61), medelst teckenigenkänning, och

2000 -11- 1 3

Huvudfoxen Kassan

20

att pussla ihop teckenuppsättningarna (50, 51) med
ledning av tecknen (60, 61) i teckenuppsättningarna (50,
51).

0011131728FAX+4640260516

Ink. i Patent- och reg.verket

2000 -11- 13

Huvudfaxen Kassan

21

SAMMANDRAG

- En anordning och ett förfarande för registrering av text genom avbildning av texten på en ljuskänslig sensor
- 5 (8) med en tvådimensionell sensoryta beskrivs.
- Anordningen är anordnad att utföra förfarandet att omvandla åtminstone två av de med sensorn registrerade bilderna (14-17, 47-49) till vars en teckenuppsättning (50, 51), innefattande ett flertal tecken (60, 61),
- 10 medelst teckenigenkänning, och att därefter pussla ihop teckenuppsättningarna (50, 51) med ledning av tecknen i teckenuppsättningarna.

1/4

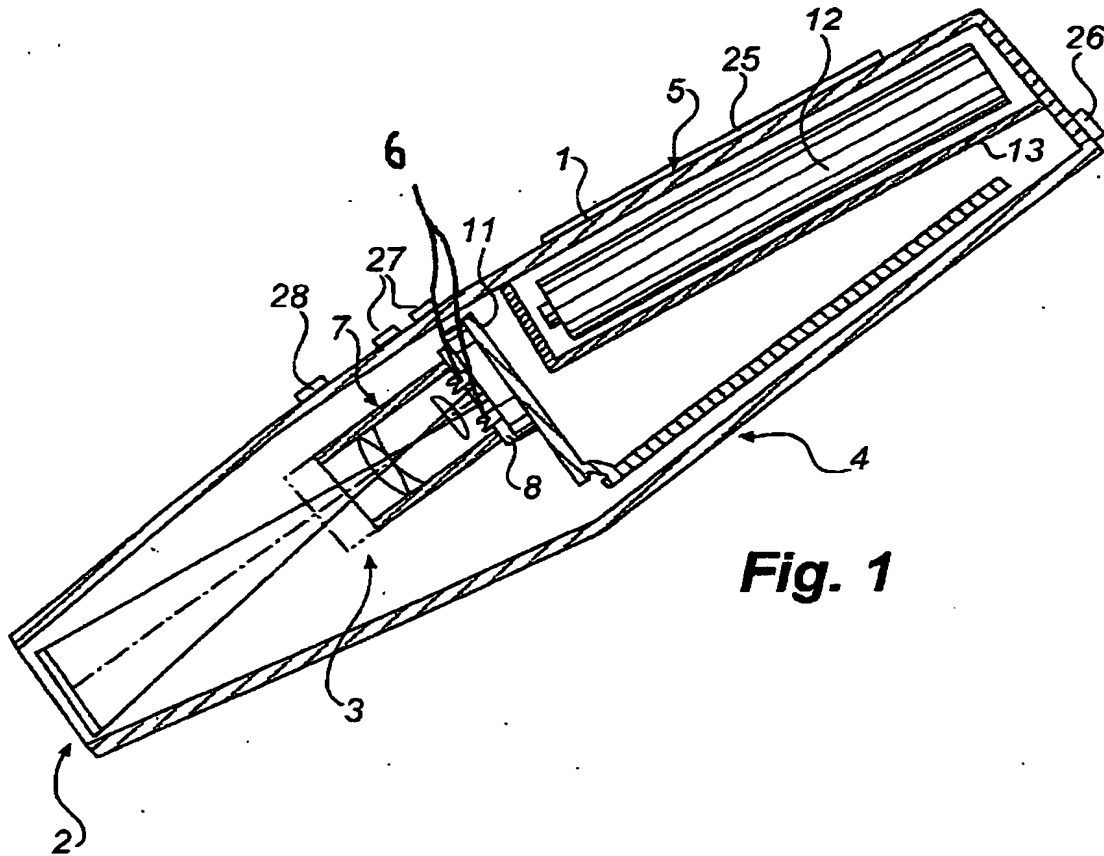


Fig. 1

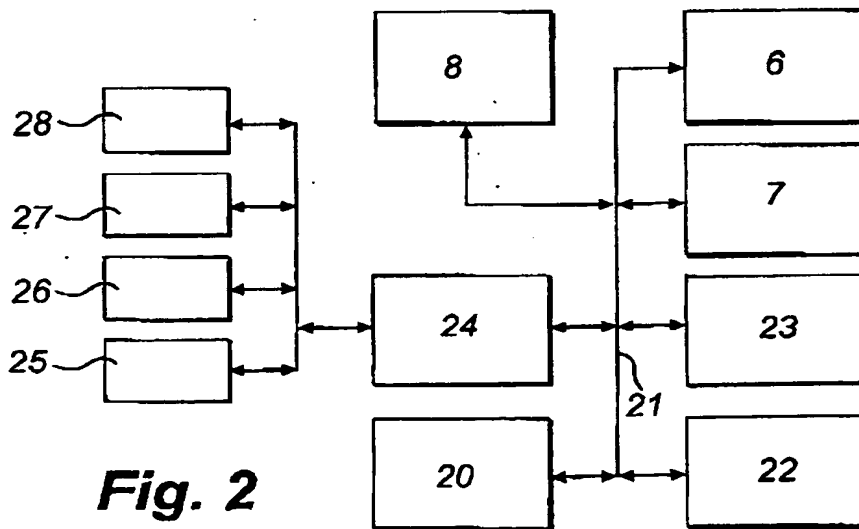


Fig. 2

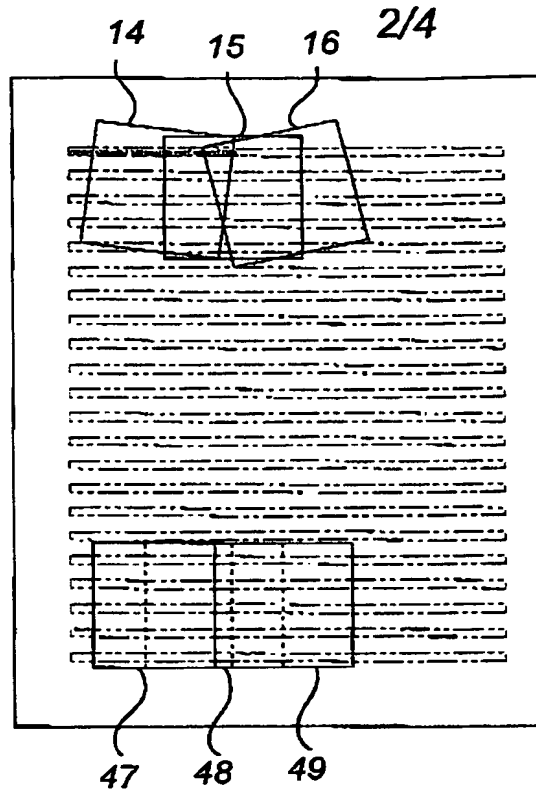


Fig. 3

52 63 64 50 53
 ner sufficiently
 by a person skilled i
 art" for this purpose
 ctitioner aware not o
 , itself and the refere
 -mmon general know
 e application. He is a
 and the c

51 54 61
 on skilled in the art
 purpose is consid
 are not only of the t
 e references therein
 al knowledge in the
 e is assumed to ha
 the capacity for

60 62

Fig. 8

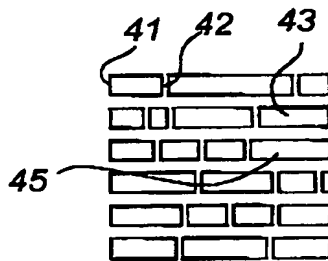


Fig. 7a

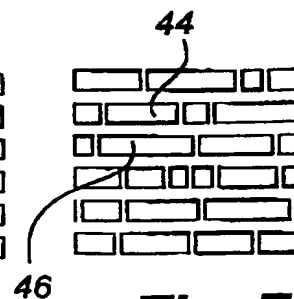


Fig. 7b

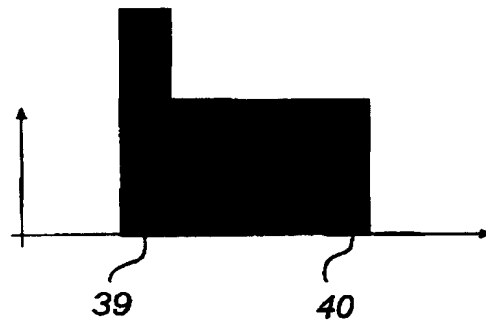
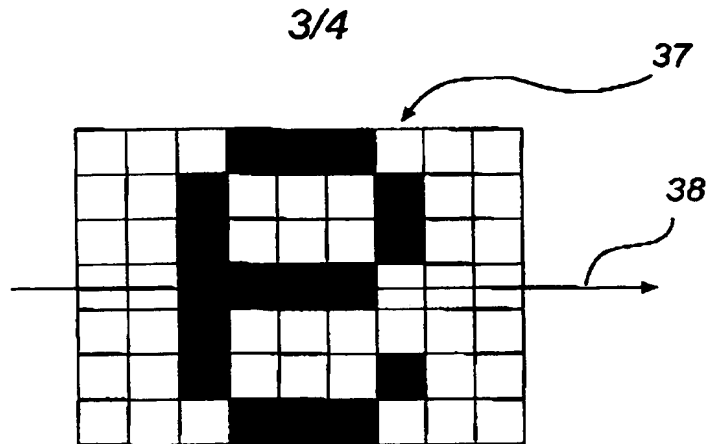


Fig. 6

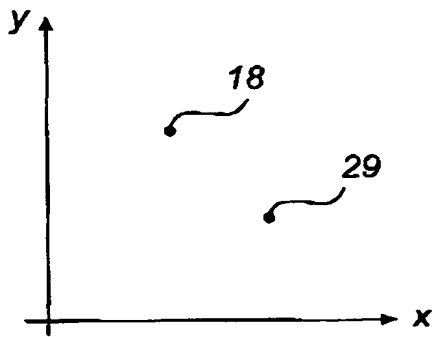


Fig. 4a

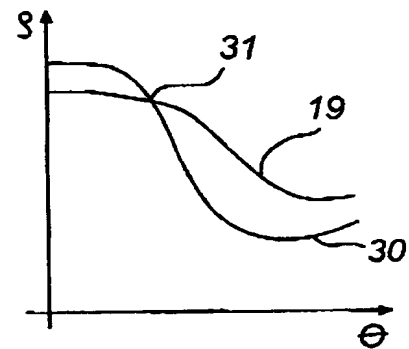


Fig. 4b

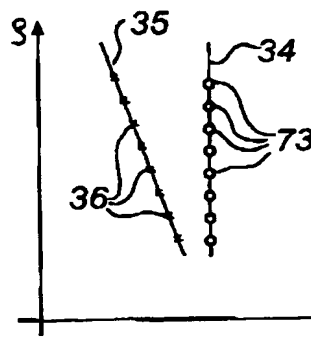


Fig. 5

4/4

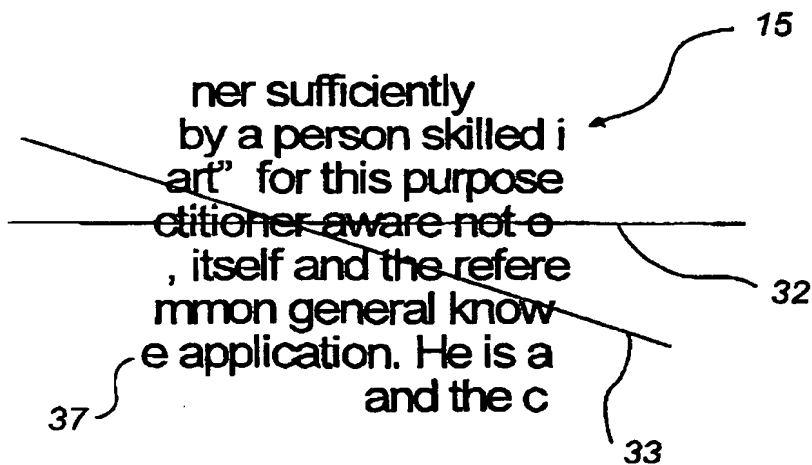


Fig. 9

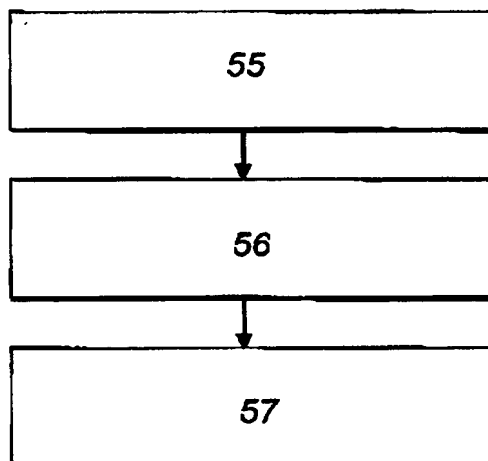


Fig. 10